

# Communicating integrated data

DATA INTEGRATION WORKSHOPS
FOR ASIA AND THE PACIFIC

NOVEMBER-DECEMBER 2020

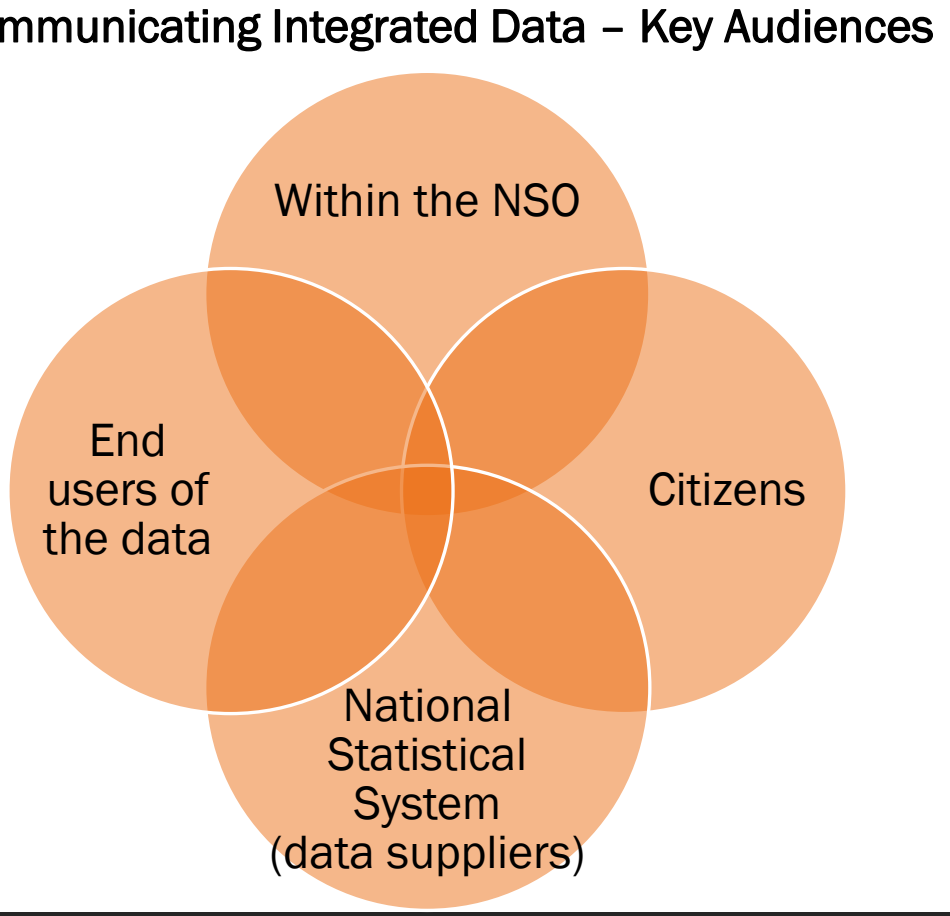
# What do we mean by ‘communication’

Communicate: to make known, impart, transmit

Two-way exchange

Communication essential throughout any data production process for:

- Understanding needs and demands for data
- Establishing data sharing agreements
- Explaining methods
- Presenting results
- Seeking and giving feedback (suppliers, users, etc.)
- Improving data quality



# What is different about communicating ‘integrated data’

### 1. Data integration relies on building relationships

- Effective communication is crucial for creating win-win partnerships with policy makers, other government agencies and private sector

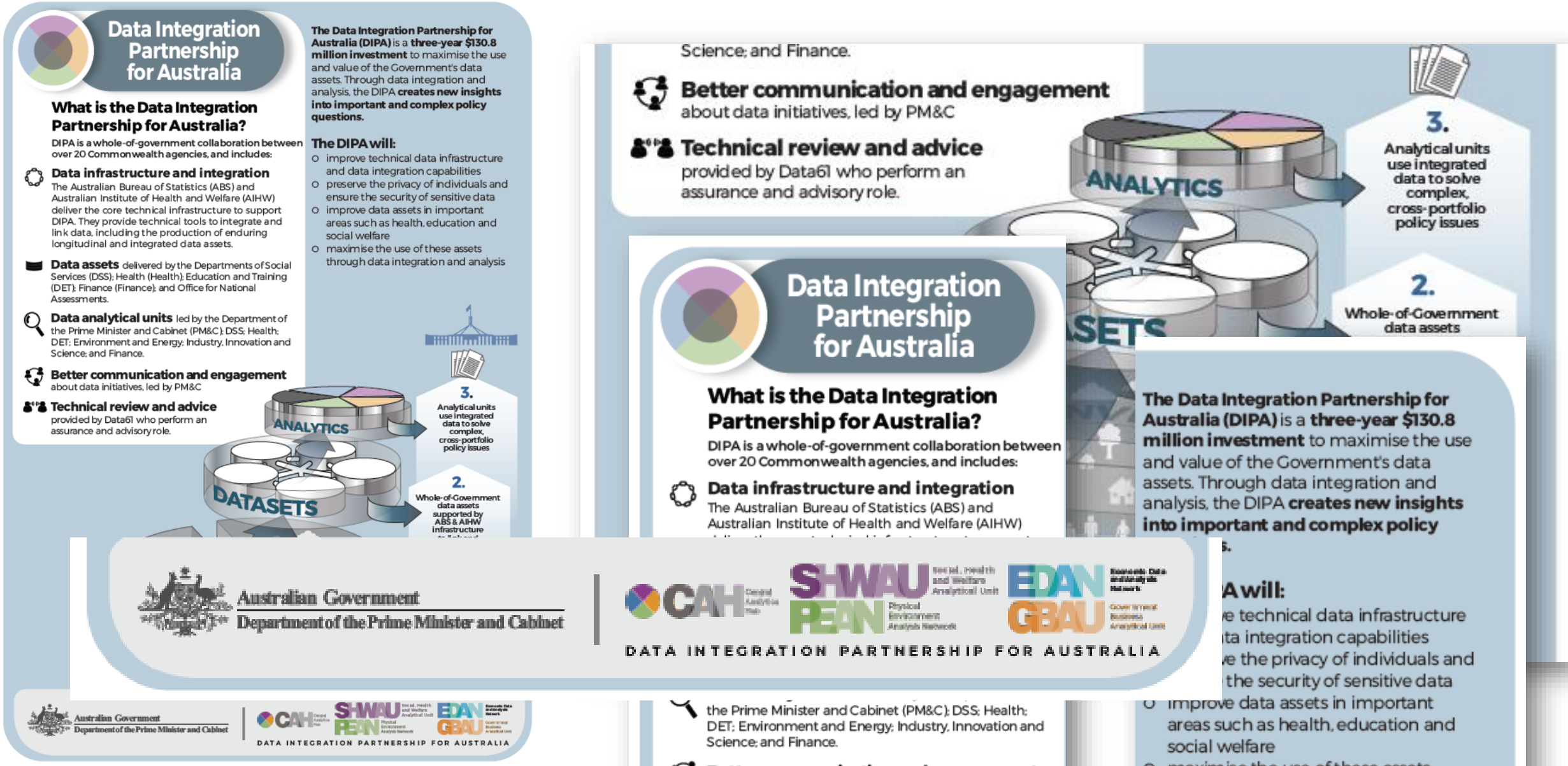
### 2. Using new, emerging, and complex methods

- Documenting approaches and writing guidelines
- Explain the value proposition of integration
- Explain methods to non-technical users

### 3. Mitigating risks to trust and credibility

- Citizens need to be informed and have their concerns about data use and confidentiality
- “NSOs should communicate that statistical confidentiality will remain in place in the statistical data store.”
- “NSOs should lead discussions to clarify the difference between concepts of privacy, security and (statistical) confidentiality”






Source: Australian Government
https://www.pmc.gov.au/public-data/data-integration-partnership-australia

# 1160.0 - ABS Confidentiality Series, Aug 2017

**LATEST ISSUE** Released at 11:30 AM (CANBERRA TIME) 23/08/2017 **First Issue**



Introduction
**What is confidentiality and why is it important**
Understanding re-identification
Managing the risk of disclosure: the Five Safes Framework
Managing the risk of disclosure: treating aggregate data
Managing the risk of disclosure: treating microdata
ABS Microdata: uses and impact on research quality
About this Release

## WHAT IS CONFIDENTIALITY AND WHY IS IT IMPORTANT?

This page contains the following:

What is confidentiality?
The obligation to maintain confidentiality
Managing data confidentiality
A contextual approach to confidentiality

### WHAT IS CONFIDENTIALITY?

Confidentiality refers to protecting the secrecy and privacy of information collected from individuals and organisations.

This means that when information is made available, it needs to be done in a way that is unlikely to allow individuals or organisations to be identified. Maintaining confidentiality is both a legal and ethical obligation, and a failure to maintain confidentiality is called a confidentiality breach, or disclosure.

This series focuses on managing the risks of two main types of breach:

- Where the identity of a person or organisation is determined using other public or privately held information about them. This is known as re-identification.
- Where a characteristic of an individual or organisation is determined without formally re-identifying them. This is known as attribute disclosure.

Feedback

Source: https://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/1160.0Main%20Features2Aug%202017?opendocument&tabname=Summary&prodno=1160.0&issue=Aug%202017&num=&view=

# New legislation

The Australian Government is working to develop better, more seamless services to the public. To do this, we need to modernise how we manage the wealth of information supplied by Australians to various government agencies.

New legislation to support better sharing of government held data builds on the Government's response to recommendations made by the Productivity Commission's Data Availability and Use Inquiry.

As part of getting these reforms right, we have engaged widely and consulted on an Issues Paper in July-August 2018, followed by a Discussion Paper in September-October 2019.

> We have taken this feedback on board; we are progressing the data reform legislation, previously known as the Data Sharing and Release legislation, as the Data Availability and Transparency legislation. When the legislation passes Parliament, it will be known as the Data Availability and Transparency Act (DATA).



You are at

Summary
We need reform
Sharing for the public interest
Streamlined, safe and secure
More than legislation
Out of scope
An evolving system

## Summary

The reforms are about streamlining the sharing of government data so we can deliver better services, better decisions and better government, for all Australians.

Under the legislation, data will only be shared for three purposes:

01 **Government services delivery** – more convenient access to the information and services people need.

02 **Government policy and programs** – planning for the future based on the best available information.

03 **Research and development** – to guide the country's academics, scientists and innovators to make our economy, environment and society healthier into the future.

Modern and streamlined legislation will:

- provide government agencies (Data Custodians) with an authorisation to share government data to accredited users such as government agencies, state and territory authorities, and non-government entities such as universities
- take a holistic risk management approach to safeguard the sharing of government data
- empower the National Data Commissioner to develop requirements and guidance to support

You are at

Summary
We need reform
Sharing for the public interest
Streamlined, safe and secure
More than legislation
Out of scope
An evolving system

Expand all +

## We need reform

## Sharing for the public interest

## Streamlined, safe and secure

We understand the need to protect your privacy and be transparent about how your data is used, the legislation will require that data is only shared where it is safe and secure. There will be transparency and accountability mechanisms put in place so that you can find out what data is being shared, with who and for what purpose, including publishing on our website:

- agreements for data shared under the legislation
- a list of agencies, businesses, individuals and organisations accredited by us
- annual reports on the operation of the system.

It will build on existing safeguards to bring a more consistent and transparent approach to how the public service shares data. This includes consistent request forms, requirements and purposes across all data sharing under the legislation.

Government agencies will apply an internationally recognised risk management approach when they consider a request for their data. If the risks cannot be appropriately managed and the safeguards are not

You are at

Summary
We need reform
Sharing for the public interest
**Streamlined, safe and secure**
More than legislation
Out of scope
An evolving system



Statistics New Zealand
TATAURANGA AOTEAROA

# Data integration manual: 2nd edition

New Zealand Government

## Contents


**List of tables and figures** ..... **5**

**1 Purpose and summary** ..... **6**
1.1 Manual's structure ..... 6

**2 Introduction to data integration** ..... **7**
2.1 Key points about data integration ..... 7
2.2 Key data integration concepts ..... 8
2.3 Data integration at Statistics NZ and elsewhere ..... 10
2.4 Key steps in a data integration project ..... 11

**3 Legal and policy considerations** ..... **12**
3.1 The Statistics Act 1975 ..... 12
3.2 The Privacy Act 1993 ..... 13
3.3 Statistics NZ's data integration policy ..... 14
3.4 Codes of practice ..... 15
3.5 Data integration business case ..... 15
3.6 Statistics NZ's confidentiality protocol ..... 17
3.7 Statistics NZ's microdata access protocols ..... 17

**4 Operational aspects of a Statistics NZ data integration project** .....
4.1 Early stages of a project .....
4.2 Other relationships .....
4.3 Obtaining and keeping external data safe .....
4.4 Documentation and quality assurance .....
4.5 Information technology considerations .....

**5 Preparing for record linkage** .....
5.1 Gathering information about source data .....
5.2 Procedure for obtaining data .....
5.3 Preparing data for record linkage .....

**6 Statistical theory of record linkage** .....
6.1 Exact linking .....
6.2 Matches and links .....
6.3 Linking files .....
6.4 Weights .....
6.5 Blocking .....
6.6 Passes .....


## 1 Purpose and summary

The *Data integration manual* 2nd edition provides a guide to data integration at Statistics New Zealand. The manual was written by Statistics NZ staff, following our involvement in several large inter-agency data integration projects.

The manual's purpose is to guide best practice and share the insights gained from our experience. The manual will assist agencies collaborating with Statistics NZ, and others interested in data integration, to understand the basic concepts, theory, and processes involved in data integration, as well as providing practical advice.

### 1.1 Manual's structure

The manual begins with an introduction that describes what data integration is and why it is carried out, and outlines the key steps involved.

Chapter 3 introduces the legal environment and Statistics NZ policy on data integration.

Chapter 4 describes operational aspects of Statistics NZ data integration projects.

The remaining chapters focus on technical aspects of the linkage itself: the data preparation needed before record linkage can be undertaken, the statistical theory of record linkage, and how to practically implement record linkage techniques.

Source:
http://archive.stats.govt.nz/~/media/Statistics/surveys-and-methods/methods/data-integration/data-manual-2ed/data-integr-manual-2edn.pdf

Stats NZ
Tatauranga Aotearoa

STATISTICS TOOLS SERVICES AND SUPPORT INTEGRATED DATA CENSUS WELLBEING INDICATORS ABOUT US

**Integrated data**

Integrated Data Infrastructure

Longitudinal Business Database

Apply to use microdata for research

If you can't find what you're looking for, contact us via live chat (bottom right of this page), email info@stats.govt.nz or call 0508 525 525.

FIND OUT MORE

**Estimated population of NZ**
5,101,400
at 30 September 2020

**Unemployment rate**
5.3%
September 2020 quarter

**Consumers price index**
+1.4%
September 2020 year

**Gross domestic product**
-12.2%
June 2020 quarter

## Statistics

POPULATION

BUSINESS

## Data tools

**COVID-19 data portal**
Data on NZ's economy since the COVID-19 outbreak.

Chat with Stats NZ now!

# Integrated Data Infrastructure

02 July 2018, 11:30am

The Integrated Data Infrastructure (IDI) is a large research database. It holds microdata about people and households.

The data is about life events, like education, income, benefits, migration, justice, and health. It comes from government agencies, Stats NZ surveys, and non-government organisations (NGOs). The data is linked together, or integrated, to form the IDI.

The IDI complements the Longitudinal Business Database (LBD), which holds linked microdata about businesses. The two databases are linked through tax data.

Researchers use the IDI to gain insight into our society and economy. The research can help answer questions about complex issues that affect New Zealanders.

See:

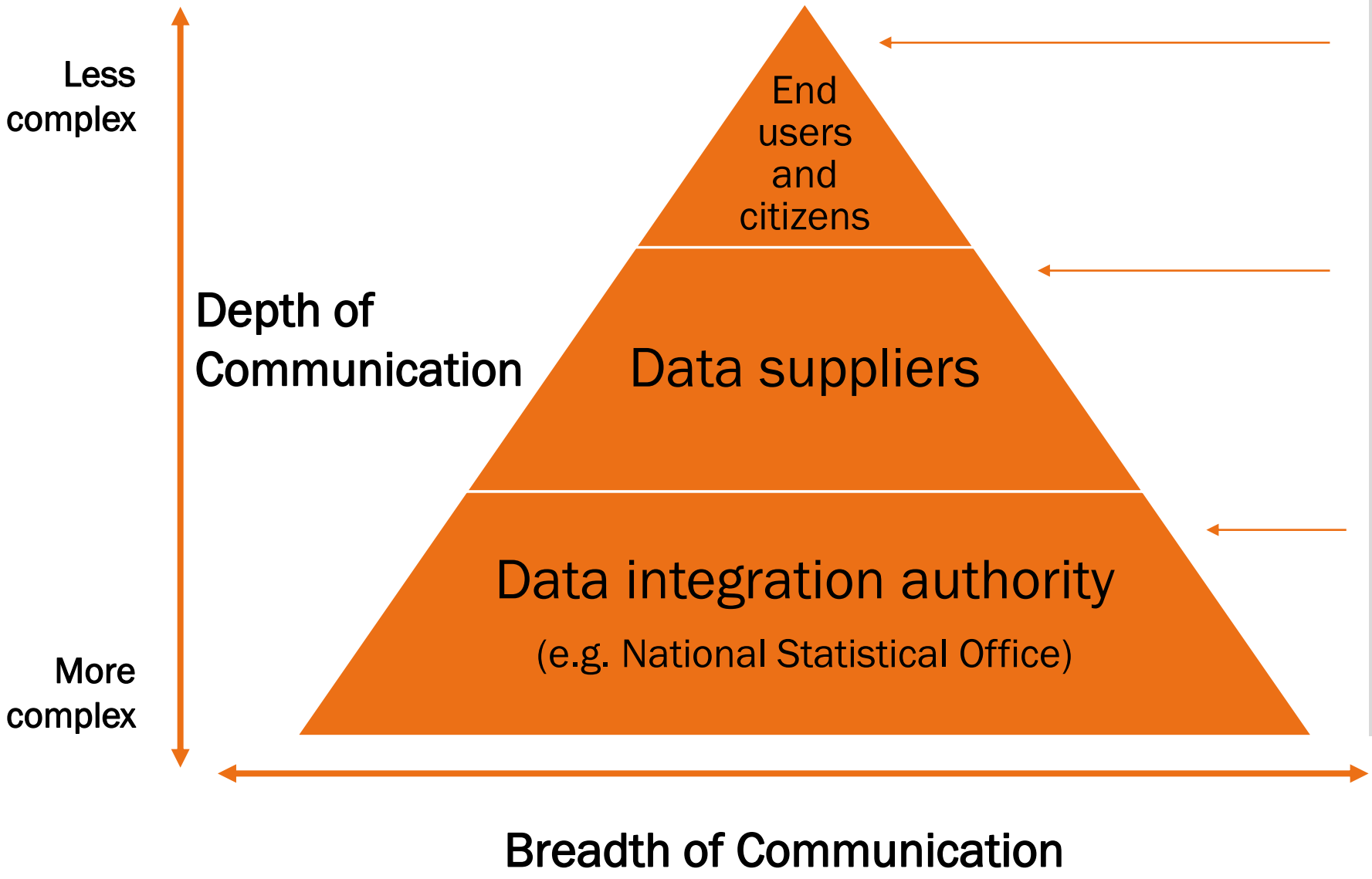
- How the IDI is being used
- IDI data does not identify individuals

Related pages

Longitudinal Business Database
02 July 2018

Chat with Stats NZ now!

# Potential Audience Communication Model




## Example

Summary information on integration methods and results

Guidance on data harmonization, management and transmission to the integration authority

Documentation on integration methods; policies for working with data suppliers; detailed analysis and presenting results for end use

# Today's task:
## Developing a checklist for communicating integrated data

What are the **considerations** and '**things to do**' for effective communication of integrated data?

Brainstorming ideas related to different target audiences and communication needs

Go to www.menti.com and enter the code:

**77 88 83 5**

